

## PATENT ABSTRACTS OF JAPAN

(11)Publication number : 2000-293685

(43)Date of publication of application : 20.10.2000

(51)Int.Cl.

G06T 7/00

(21)Application number : 11-099092

(71)Applicant : TOYOTA MOTOR CORP

(22)Date of filing : 06.04.1999

(72)Inventor : HATAKEYAMA YOSHIYUKI

OKAMURA HIROKI

NAKAMURA SHINJI

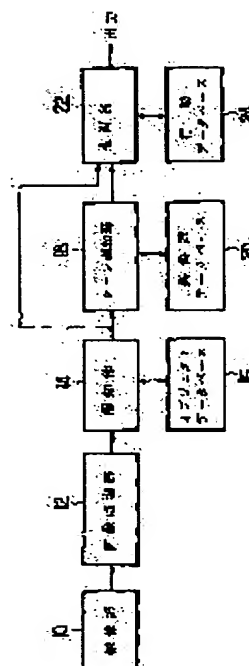
HARA YOSHIKO

## (54) SCENE RECOGNIZING DEVICE

## (57)Abstract:

PROBLEM TO BE SOLVED: To automatically recognize the scene of a picture and the action intention of a person in the picture.

SOLUTION: A picture processing part 12 extracts an object in a picture obtained by a photographing part 10. A recognizing part 14 decides the attribute of the extracted object (that is, whether or not the object is a person, and the action of a person when the object is a person, and the designation of a body when the object is a body). A scene recognizing part 18 selects the most certain scene based on the attribute of the object from among scene candidates. An estimating part 22 specifies the most matched action intention of the person from among the plural possible action intentions based on the action attribute of the person object in the recognized scene.



(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号  
特開2000-293685  
(P2000-293685A)

(43) 公開日 平成12年10月20日 (2000. 10. 20)

(51) Int.Cl.<sup>7</sup>  
G 0 6 T 7/00

識別記号

F. I  
G 0 6 F 15/70

キーワード (参考)

4 6 0 Z 5 L 0 9 6

審査請求 未請求 請求項の数 5 O L (全 14 頁)

(21) 出願番号 特願平11-99092

(22) 出願日 平成11年4月6日 (1999. 4. 6)

(71) 出願人 000003207

トヨタ自動車株式会社

愛知県豊田市トヨタ町1番地

(72) 発明者 島山 善幸

愛知県豊田市トヨタ町1番地 トヨタ自動車株式会社内

(72) 発明者 岡村 宏樹

愛知県豊田市トヨタ町1番地 トヨタ自動車株式会社内

(74) 代理人 100075258

弁理士 吉田 研二 (外2名)

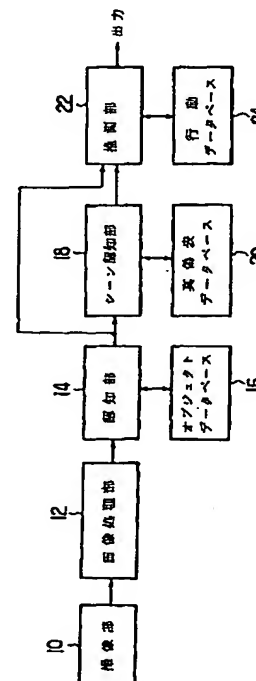
最終頁に続く

(54) 【発明の名称】 シーン認識装置

(57) 【要約】

【課題】 画像のシーン、さらには画像内の人物の行動意図を自動認識する。

【解決手段】 画像処理部12は、撮像部10で取得した画像内のオブジェクトを抽出する。認知部14は、抽出したオブジェクトの属性（人物か否か、人物であればその行動、物体であればその名称）を決定する。シーン認知部18は、シーン候補の中からオブジェクトの属性に基づいて最も確からしいシーンを選択する。推測部22は、認知されたシーンにおける人物オブジェクトの行動属性に基づき、複数の可能な行動意図から最も合致する行動意図を特定する。



**【特許請求の範囲】**

**【請求項1】** 画像内のオブジェクトを抽出する抽出手段と、  
抽出された前記オブジェクトの属性を認識する属性認識手段と、  
前記オブジェクトの属性に基づいて前記画像のシーンを認識するシーン認識手段と、  
を有することを特徴とするシーン認識装置。

**【請求項2】** 請求項1記載の装置において、さらに、  
前記シーン認識装置で認識されたシーン及び前記オブジェクト内の人物オブジェクトの属性に基づいて前記人物の行動を認識する行動認識手段を有することを特徴とするシーン認識装置。

**【請求項3】** 請求項2記載の装置において、  
前記行動認識手段は、前記シーン認識手段で認識されたシーンに対する複数の可能な行動の中から、前記人物オブジェクトの属性に合致する行動を選択し、選択した行動から前記オブジェクト内の物体オブジェクトに基づいて人物の行動意図を特定することを特徴とするシーン認識装置。

**【請求項4】** 請求項1～3のいずれかに記載の装置において、  
前記シーン認識手段は、シーンとオブジェクトの可能な組み合わせの中から、抽出された前記オブジェクトを有するシーンを選択することによりシーンを認識することを特徴とするシーン認識装置。

**【請求項5】** 請求項1～4のいずれかに記載の装置において、  
前記シーン認識手段は、シーン候補が複数存在する場合には、前記シーン候補と前記オブジェクト内の人物オブジェクトの属性に基づいて選択することを特徴とするシーン認識装置。

**【発明の詳細な説明】****【0001】**

**【発明の属する技術分野】** 本発明はシーン認識装置、特に画像内のシーンを自動認識する技術に関する。

**【0002】**

**【従来の技術】** 日常生活では、人間の行動の意図を推測しなければならない場面が少なくない。例えば、車両を運転している場合には、歩行者が横断歩道の端に立っていれば横断歩道を渡りたいという意思表示であると推測されることから、歩行者を視認した場合にはその意図を迅速に把握して車両を停止させる必要がある。一方、同じ歩行者が立っていても、近くに横断歩道が存在しない場合には、必ずしも道路を横断する意図があるとは言えない。

**【0003】** また、レストランで顧客が手を挙げた場合、店員は顧客が手を挙げた意図を周囲状況から判断して適切な対応をとる必要がある。テーブル上に伝票がある場合には勘定を支払うという意図があり、伝票がない

場合には追加注文をする意図があると判断する等である。

**【0004】** さらに、ビルの警備では、警備員は監視カメラの画像内の人物の行動から不審人物であるか否かを正確に判断する必要がある。

**【0005】**

**【発明が解決しようとする課題】** しかしながら、推測結果は推測者個人の経験や精神状態、体調などの条件に左右されることがあり、常に一定の精度で行動意図を推測することは困難である問題がある。例えば、経験が少ない店員では顧客の行動の意図を正確に把握することは困難であり、また24時間の警備を行う場合には経験豊かな警備員でも特に深夜から早朝にかけて判断力が低下することもあり得る。

**【0006】** 本発明は、上記従来技術の有する課題に鑑みなされたものであり、その目的は、画像の意味するシーンを機械的に自動認識でき、さらにそのシーンにおける人物の行動意図を機械的に自動認識することができる装置を提供することにある。

**【0007】**

**【課題を解決するための手段】** 上記目的を達成するために、本発明は、画像内のオブジェクトを抽出する抽出手段と、抽出された前記オブジェクトの属性を認識する属性認識手段と、前記オブジェクトの属性に基づいて前記画像のシーンを認識するシーン認識手段とを有することを特徴とする。オブジェクトの属性とは、具体的にはオブジェクトが人物であるか物体であるか、そして、人物である場合にはどのような行動をしているか、物体である場合にはその物体は何かを示すものであり、これらの組み合わせに基づいて画像がどんなシーンであるかを認識することができる。

**【0008】** ここで、前記シーン認識装置で認識されたシーン及び前記オブジェクト内の人物オブジェクトの属性に基づいて前記人物の行動を認識する行動認識手段をさらに有することもできる。これにより、人物の行動意図に応じた対応をとることが可能となる。

**【0009】** 前記行動認識手段は、前記シーン認識手段で認識されたシーンに対する複数の可能な行動の中から、前記人物オブジェクトの属性に合致する行動を選択し、選択した行動から前記オブジェクト内の物体オブジェクトに基づいて人物の行動意図を特定する。

**【0010】** 前記シーン認識手段は、シーンとオブジェクトの可能な組み合わせの中から、抽出された前記オブジェクトを有するシーンを選択することによりシーンを認識する。

**【0011】** また、前記シーン認識手段は、シーン候補が複数存在する場合には、前記シーン候補と前記オブジェクト内の人物オブジェクトの属性に基づいてシーンを選択する。

**【0012】**

【発明の実施の形態】以下、図面に基づき本発明の実施形態について説明する。

【0013】図1には、本実施形態の構成ブロック図が示されている。本実施形態の装置は、撮像部10、画像処理部12、認知部14、オブジェクトデータベース16、シーン認知部18、真偽表データベース20、推測部22及び行動データベース24を含んで構成されている。

【0014】撮像部10は、CCDカメラなどで構成され、人物を含む画像を取得する。取得した画像は画像処理部12に供給される。もちろん、他の撮像部で得られた画像を記録メディアに記録し、記録メディアから読み出して画像処理部12に供給してもよい。撮像部10は必須の部材ではない。

【0015】画像処理部12は、入力した画像内に存在する人物像や物体像（人物像と物体像をまとめてオブジェクトと称する）を抽出して認知部14に供給する。オブジェクトの抽出は、例えばオブジェクト間の意図の違いや輝度の違い（エッジ）に基づいて行うことができる。抽出したオブジェクトは人物と物体の区別なく、全て認知部14に供給される。

【0016】認知部14は、オブジェクトデータベース16に記憶されている基準オブジェクトと入力されたオブジェクトとを対比し（テンプレートマッチング）、入力されたオブジェクトの属性、すなわち入力オブジェクトが人物オブジェクトか物体オブジェクトか、人物オブジェクトの場合にはどのような動作を行っているのか、物体オブジェクトの場合にはその物体の名称はなにか、を決定する。オブジェクトの属性を正確に決定するためには、オブジェクトデータベース16に予め記憶する基準オブジェクトは可能な限り多種多様であることが好ましい。基準オブジェクトの例としては、人物に関しては、立っている人物、座っている人物、腕を組んでいる人物、手を挙げている人物、歌っている人物、踊っている人物、歩いている人物、走っている人物等々であり、物体に関しては、机、テーブル、いす、本、コンピュータ、棒、はし、飲食器、信号機等々である。認知したオブジェクトは、その属性に基づいて人間群ファイルと物体群ファイルに分けてシーン認知部18に供給される。

【0017】シーン認知部18は、人間群ファイルと物体群ファイルに基づき、撮像部10から供給された画像がどんなシーンの画像であるかを認知する。シーンの認知は、真偽表データベース20に記憶されている真偽表に基づき行われる。真偽表データベース20は、可能なシーンそれぞれについて種々のオブジェクトが存在するか否かを○か×かで示したもので（そのシーンにそのオブジェクトが存在する場合には○、存在しない場合には×）、例えばレストランのシーンにおいては皿類やフォークというオブジェクトは存在し得るので○、信号機は存在し得ないので×、走っている人物も通常は存在しな

いので×等である。シーン認知部18は、このような真偽表に基づいて、入力したオブジェクトが最も数多く存在し得るシーンを抽出し、画像のシーンであると認知する。認知したシーンは、推測部22に供給される。

【0018】推測部22は、認知したシーン及び認知部14からのオブジェクトの属性に基づいて、認識したシーンにおける人物の行動意図を推測して出力する。行動意図は、行動データベース24に記憶されている、各シーンにおいて考えられる行動の意図から最も合致する意図を選択することにより推測される。

【0019】画像処理部12、認知部14、シーン認知部18、推測部22はコンピュータで構成することができ、オブジェクトデータベース16や真偽表データベース20、行動データベース24は磁気テープや磁気ディスク、光ディスクなどの大容量メモリで構成することができる。

【0020】以下、認知部14、シーン認知部18、推測部22の処理について、具体的な例を挙げて説明する。

【0021】＜レストランのテーブルに座った人物が挙手をしている画像が入力された場合＞このような画像が入力された場合、画像処理部12では、人物オブジェクトと物体オブジェクトを抽出して認知部14に供給する。認知部14では、これらのオブジェクトの属性を判別して人間群ファイルと物体群ファイルに分類する。

【0022】図2には、認知部14の処理フローチャートが示されている。まず、画像処理部12から抽出したオブジェクトのファイルを入力し（S101）、順次オブジェクトデータベース16に記憶されている基準オブジェクトと対比することにより人物オブジェクトか物体オブジェクトかを判別する（S102）。そして、入力オブジェクトが人物オブジェクトである場合には、さらに基準オブジェクトとの対比からその人物がどういう行動を行っているかを判別する（S103）。具体的には、その人物が座っているか、立っているか、挙手しているか、うつむいているか等である。そして、人物オブジェクトである場合には、その行動属性とともに人間群ファイルにまとめる（S104）。

【0023】一方、入力オブジェクトが物体オブジェクトである場合には、さらにその物体の名称を基準オブジェクトとの対比により特定する（S105）。具体的には、その物体がテーブルであるか、信号機であるか、時計であるか、等である。オブジェクトの属性の決定精度は、基準オブジェクトの数に依存するので、オブジェクトデータベースのデータ量は十分確保する必要がある。そして、物体オブジェクトは、その名称属性とともに物体群ファイルにまとめる（S106）。以上のようにしてオブジェクトの属性が認識されると、その結果をシーン認知部18に出力する（S107）。

【0024】図3には、シーン認知部18の処理フロー

チャートが示されている。まず、人間群ファイルと物体群ファイルを入力すると(S201)、真偽表データベース20内の真偽表を用いて画像のシーン候補を抽出する(S202)。図3には、真偽表の一例が示されており、行はオブジェクト、列はシーンを示している。そして、各シーンにおいてそのオブジェクトが存在し得る場合には真(O)、存在し得ない場合には偽(X)が付されている。例えば、レストランのシーンでは、皿類、フォーク、食べるという行動、挙手という行動は存在し得るのでO、信号機や走るという行動、スライド映写機は存在し得ないのでX、横断歩道のシーンでは信号機や走るという行動、挙手という行動は存在し得るのでO、逆に皿類やフォーク、食べるという行動は存在し得ないのでX、会議室のシーンでは皿類やスライド映写機、挙手という行動は存在し得るのでO、フォークや信号機、走るという行動、食べるという行動は存在し得ないのでX等である。このような真偽表を用いてそれぞれのシーンに対して入力したオブジェクトの真偽を判定し、真

(O)の数の多いシーン候補を抽出していく。シーン候補を高精度に抽出するためには、真偽表におけるシーンの種類及びオブジェクトの種類を可能な限り多くしておくことが好適である。真偽表のオブジェクトは基準オブジェクトと同一とすることができ、真偽表のシーンは認知を希望する全てのシーンをカバーするのが好ましい。

【0025】抽出したシーン候補が単一の場合には、そのシーン候補をもって入力画像のシーンであると認定することが可能であるが、シーン候補が複数存在する場合、つまり1つでも真(O)があるシーンが複数存在する場合には、これらのいずれであるかを特定する必要がある。

【0026】そこで、真偽表を用いて真偽を判定した後、シーン候補を真(O)の数の多い順に並べ替える(S203)。本画像の場合、オブジェクトとして皿類やフォーク、座る、挙手が抽出されるとすると、レストランのシーン、会議室のシーン、横断歩道の順に真(O)の数が多いため、1. レストラン、2. 会議室、3. 横断歩道・・・というように順位付ける。

【0027】そして、人間群ファイルから人物オブジェクトの属性を入力し、順位付けたシーン候補のそれぞれに対して、そのシーン内で人物オブジェクトの行動属性が適当か否か、つまり、そのシーンにおいてそのような行動をとることがあり得るか否かを判定する(S204)。例えば、S202で抽出したシーン候補として横断歩道があるが、人物オブジェクトの属性が座っていて挙手をしているものである場合、横断歩道では座って挙手することは通常あり得ないので横断歩道をシーン候補から削除する。この判定は、S202で用いた真偽表を再度用いて行うことができる。以上の処理を順位の上位から順次行うことで、最終的に1つのシーンを選択して(複数残った場合には最も上位のシーン)画像のシーン

であると認知する。本画像の場合、最上位のレストランのシーンが選択される。認知されたシーンは、推測部22に出力する(S205)。

【0028】図4には、推測部22の処理フローチャートが示されている。まず、シーン認知部18からの情報、すなわち認知したシーンを入力し(S301)、また、認知部14から人物オブジェクトの中で行動意図を推測すべき人物オブジェクトの行動属性(本画像では、挙手という行動属性)も入力して、認知したシーンにおけるその人物の行動意図を推測する(S302)。行動意図の推測には、行動データベース24に記憶されている推測カードを用いる。推測カードは、シーンにおける可能な行動毎に分類されてデータベース化されており、認知したシーン及び人物オブジェクトの行動属性に合致するカード(この場合、レストランにおける挙手というカード)を抽出する。すなわち、行動データベース内に記憶されている、シーン毎、及びシーンにおける可能な行動毎に階層的に分類されている推測カードから、認知したシーン及び人物オブジェクトの行動属性を指数として、最適の推測カードを抽出するのである。各推測カードは、図に示すように、物品の有無とその組み合わせに対応する行動意図候補がリスト化されて構成されている。カード中のOはその物品が存在することを示し、Xは存在しないことを示している。そこで、抽出したカード内において人物オブジェクトの回りに存在する物体オブジェクトの有無が、カード内の有無のパターンと一致するものを最終的な行動意図として推測する(S303)。例えば、挙手をした人物の回りに皿オブジェクトが存在し、メニューオブジェクトもレシート(伝票)オブジェクトも存在しない場合には、推測カードの中から「1. 追加注文、2. 料理催促」という行動意図候補を特定する。これは、挙手をしている人物が、追加注文を要求しているか、あるいは料理を催促しているかのいずれかであることを意味する。また、挙手をした人物の回りに皿オブジェクトとメニューオブジェクトとレシートオブジェクトが存在する場合には、推測カードの中から「1. 会計、2. 水催促」という行動意図候補を特定して出力する。これは、挙手をしている人物が、会計を要求しているか、水を催促しているかのいずれかであることを意味している。そして、特定した行動意図を外部に出力する(S304)。

【0029】行動意図の推測精度は、推測カードの量及び各推測カードにおける可能な行動意図の量に依存して決定されるので、可能な限り多数の推測カードをデータベース化するのが好適である。以上のようにして、入力画像からその画像のシーン及び人物の行動意図を自動的に認識することができる。

【0030】<オフィスのドアの前で棒を持った人物が存在する画像が入力された場合>このような画像が入力された場合、画像処理部12では、人物オブジェクトと

物体オブジェクトを抽出して認知部14に供給する。認知部14では、これらのオブジェクトの属性を判別して人間群ファイルと物体群ファイルに分類する。人物群ファイルとしては、立って両手を挙手している人物、物体群ファイルとしては、ドア、棒、ホワイトボード等である。

【0031】図5には、シーン認知部18の処理フローチャートが示されている。まず、人間群ファイルと物体群ファイルを入力すると(S401)、真偽表データベース20内の真偽表を用いて画像のシーン候補を抽出する(S402)。図5には、真偽表の一例が示されており、行はオブジェクト、列はシーンを示している。そして、各シーンにおいてそのオブジェクトが存在し得る場合には真(O)、存在し得ない場合には偽(X)が付されている。例えば、オフィスのシーンでは、ドア、ホワイトボード、棒、立位という行動は存在し得るのでO、食するという行動は通常存在しないのでX、廊下のシーンでは立位という行動は存在し得るのでO、逆にドアや棒は存在し得ないのでX、給湯室のシーンではドアや食するという行動は存在し得るのでO、ホワイトボードや両手挙手という行動は存在し得ないのでX等である。このような真偽表を用いてそれぞれのシーンに対して入力したオブジェクトの真偽を判定し、真(O)の数の多いシーン候補を抽出していく。

【0032】抽出したシーン候補が単一の場合には、そのシーン候補をもって入力画像のシーンであると認定することが可能であるが、シーン候補が複数存在する場合、つまり1つでも真(O)があるシーンが複数存在する場合には、これらのいずれであるかを特定するため、シーン候補を真(O)の数の多い順に並べ替える(S403)。本画像の場合、オブジェクトとしてドアやホワイトボード、棒が抽出されるとすると、オフィス、給湯室、廊下の順に真(O)の数が多いため、この順に順位付ける。

【0033】そして、人間群ファイルから人物オブジェクトの属性を入力し、順位付けたシーン候補のそれぞれに対して、そのシーン内で人物オブジェクトの行動属性が適当か否か、つまり、そのシーンにおいてそのような行動をとることがあり得るか否かを判定する(S404)。例えば、S402で抽出したシーン候補として給湯室があるが、人物オブジェクトの属性が立っていて両手挙手をしているものである場合、給湯室ではこのような行動は通常あり得ないので給湯室をシーン候補から削除する。この判定は、S402で用いた真偽表を再度用いて行うことができる。以上の処理を順位の上位から順次行うことで、最終的に1つのシーンを選択して(複数残った場合には最も上位のシーン)画像のシーンであると認知する。本画像の場合、最上位のオフィスのシーンが選択される。認知されたシーンは、推測部22に出力する(S405)。

【0034】図6には、推測部22の処理フローチャートが示されている。まず、シーン認知部18からの情報、すなわち認知したシーン(オフィス)を入力し(S501)、また、認知部14から人物オブジェクトの中で行動意図を推測すべき人物オブジェクトの行動属性(本画像では、両手挙手という行動属性)も入力して、認知したシーンにおけるその人物の行動意図を推測する(S502)。行動意図の推測には、行動データベース24に記憶されている推測カードを用いる。推測カードは、シーンにおける可能な行動毎に分類されてデータベース化されており、認知したシーン及び人物オブジェクトの行動属性に合致するカード(この場合、オフィスにおける両手挙手というカード)を抽出する。さらに、抽出したカード内において人物オブジェクトの回りに存在する物体オブジェクトの有無が、カード内の有無のパターンと一致するものを最終的な行動意図として推測する(S503)。例えば、両手挙手をした人物の回りに鍵オブジェクトが存在し、IDカードも存在する場合には、推測カードの中から「ドア開閉」という行動意図候補を選択する。また、両手挙手をした人物の回りに鍵オブジェクトもIDカードオブジェクトも存在せず、棒オブジェクトが存在する場合には、推測カードの中から「ドア破壊」という行動意図候補を特定して出力する。本画像の場合、「ドア破壊」が特定されることになる。そして、特定した行動意図を外部に出力する(S504)。

【0035】<厨房の中でコーヒーに塩を入れている画像を入力した場合>このような画像が入力された場合、画像処理部12では、人物オブジェクトと物体オブジェクトを抽出して認知部14に供給する。認知部14では、これらのオブジェクトの属性を判別して人間群ファイルと物体群ファイルに分類する。人物群ファイルとしては、座って保持している人物、物体群ファイルとしては、コーヒー(コーヒーの容器)、ポット、カップ、塩等である。

【0036】図7には、シーン認知部18の処理フローチャートが示されている。まず、人間群ファイルと物体群ファイルを入力すると(S601)、真偽表データベース20内の真偽表を用いて画像のシーン候補を抽出する(S602)。図7には、真偽表の一例が示されており、行はオブジェクト、列はシーンを示している。そして、各シーンにおいてそのオブジェクトが存在し得る場合には真(O)、存在し得ない場合には偽(X)が付されている。例えば、厨房のシーンでは、コーヒー、ポット、座位という行動、食するという行動は存在し得るのでO、信号機や歌うという行動は通常存在しないのでX、交差点のシーンでは信号機は存在し得るのでO、逆にコーヒーや座位という行動は存在し得ないのでX、浴室のシーンでは座位や歌うという行動は存在し得るのでO、コーヒーや食するという行動は存在し得ないのでX

等である。このような真偽表を用いてそれぞれのシーンに対して入力したオブジェクトの真偽を判定し、真（○）の数の多いシーン候補を抽出していく。

【0037】抽出したシーン候補が単一の場合には、そのシーン候補をもって入力画像のシーンであると認定することが可能であるが、シーン候補が複数存在する場合、つまり1つでも真（○）があるシーンが複数存在する場合には、これらのいずれであるかを特定するため、シーン候補を真（○）の数の多い順に並べ替える（S603）。本画像の場合、オブジェクトとしてコーヒーやポット、座位、物体保持が抽出されるとすると、厨房、浴室、交差点の順に真（○）の数が多いため、この順に順位付ける。

【0038】そして、人間群ファイルから人物オブジェクトの属性を入力し、順位付けたシーン候補のそれぞれに対して、そのシーン内で人物オブジェクトの行動属性が適当か否か、つまり、そのシーンにおいてそのような行動をとることがあり得るか否かを判定する（S604）。例えば、S602で抽出したシーン候補として交差点があるが、人物オブジェクトの属性が座っている（座位）ものである場合、交差点ではこのような行動は通常あり得ないので交差点をシーン候補から削除する。この判定は、S602で用いた真偽表を再度用いて行うことができる。以上の処理を順位の上位から順次行うことで、最終的に1つのシーンを選択して（複数残った場合には最も上位のシーン）画像のシーンであると認知する。本画像の場合、最上位の厨房のシーンが選択される。認知されたシーンは、推測部22に出力する（S605）。

【0039】図8には、推測部22の処理フローチャートが示されている。まず、シーン認知部18からの情報、すなわち認知したシーン（厨房）を入力し（S701）、また、認知部14から人物オブジェクトの中で行動意図を推測すべき人物オブジェクトの行動属性（本画像では、座位で物体を保持している行動属性）も入力して、認知したシーンにおけるその人物の行動意図を推測する（S702）。行動意図の推測には、行動データベース24に記憶されている推測カードを用いる。推測カードは、シーンにおける可能な行動毎に分類されてデータベース化されており、認知したシーン及び人物オブジェクトの行動属性に合致するカード（この場合、厨房における座位で物体保持というカード）を抽出する。なお、カードを抽出する際に、保持している物体に応じて抽出するのが好適である。そして、抽出したカード内において人物オブジェクトの回りに存在する物体オブジェクトの有無が、カード内の有無のパターンと一致するものを最終的な行動意図として推測する（S703）。例えば、座って塩を保持している人物の回りにコーヒーオブジェクトとスプーンオブジェクトが存在する場合に

は、推測カードの中から「コーヒーに塩を入れようとしている」という行動意図候補を選択する。また、座って塩を保持している人物の回りにコーヒーオブジェクトが存在せず、スプーンオブジェクトやカップオブジェクトが存在する場合には、推測カードの中から「コーヒーカップに塩を入れようとしている」という行動意図候補を特定して出力する。本画像の場合、「コーヒーに塩を入れようとしている」が特定されることになる。そして、特定した行動意図を外部に出力する（S704）。

【0040】以上、本発明の実施形態について説明したが、本実施形態の装置は任意の装置ないしシステムと組み合わせることが可能である。例えば、本装置を車両に適用した場合、歩行者の行動意図を自動的に推測して自車と衝突の可能性がある場合には危険信号を発したり、ブレーキを作動させるなどの態様が可能である。また、本装置をレストランでのサービスに利用した場合、客の行動意図を自動推測して客のもとに速やかに店員を配置することも可能となる。また、不審者の検出を目的とした警備システム、あるいは調理講習を目的とした調理教示システムへ利用することも可能となる。また、本装置はシーンを自動的に認識するので、そのシーンに特有の初期設定を行うことなく人物の行動意図を推測できるという汎用性も有する。

【0041】

【発明の効果】以上説明したように、本発明によれば、画像の意味するシーンを機械的に自動認識でき、さらにそのシーンにおける人物の行動意図を機械的に自動認識することができる。

【図面の簡単な説明】

【図1】 本発明の実施形態の構成ブロック図である。

【図2】 図1における認知部の処理フローチャートである。

【図3】 図1におけるシーン認知部の処理フローチャートである。

【図4】 図1における推測部の処理フローチャートである。

【図5】 図1におけるシーン認知部の他の処理フローチャートである。

【図6】 図1における推測部の他の処理フローチャートである。

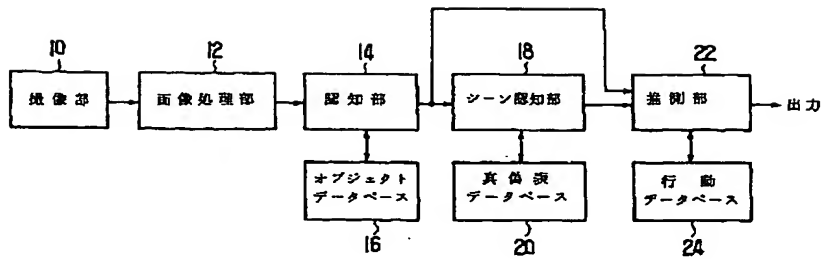
【図7】 図1におけるシーン認知部のさらに他の処理フローチャートである。

【図8】 図1における推測部のさらに他の処理フローチャートである。

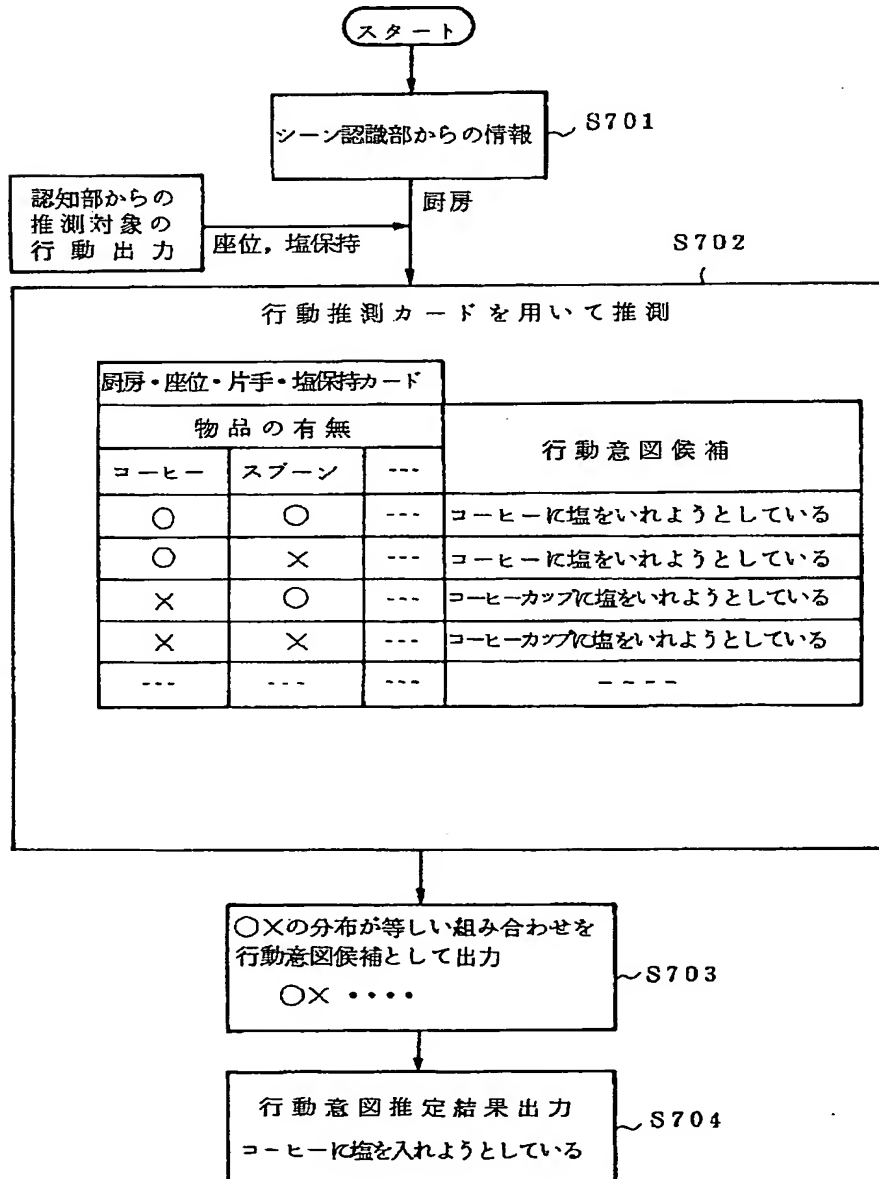
【符号の説明】

10 撮像部、12 画像処理部、14 認知部、16 オブジェクトデータベース、18 シーン認知部、20 真偽表データベース、22 推測部、24 行動データベース。

【図1】

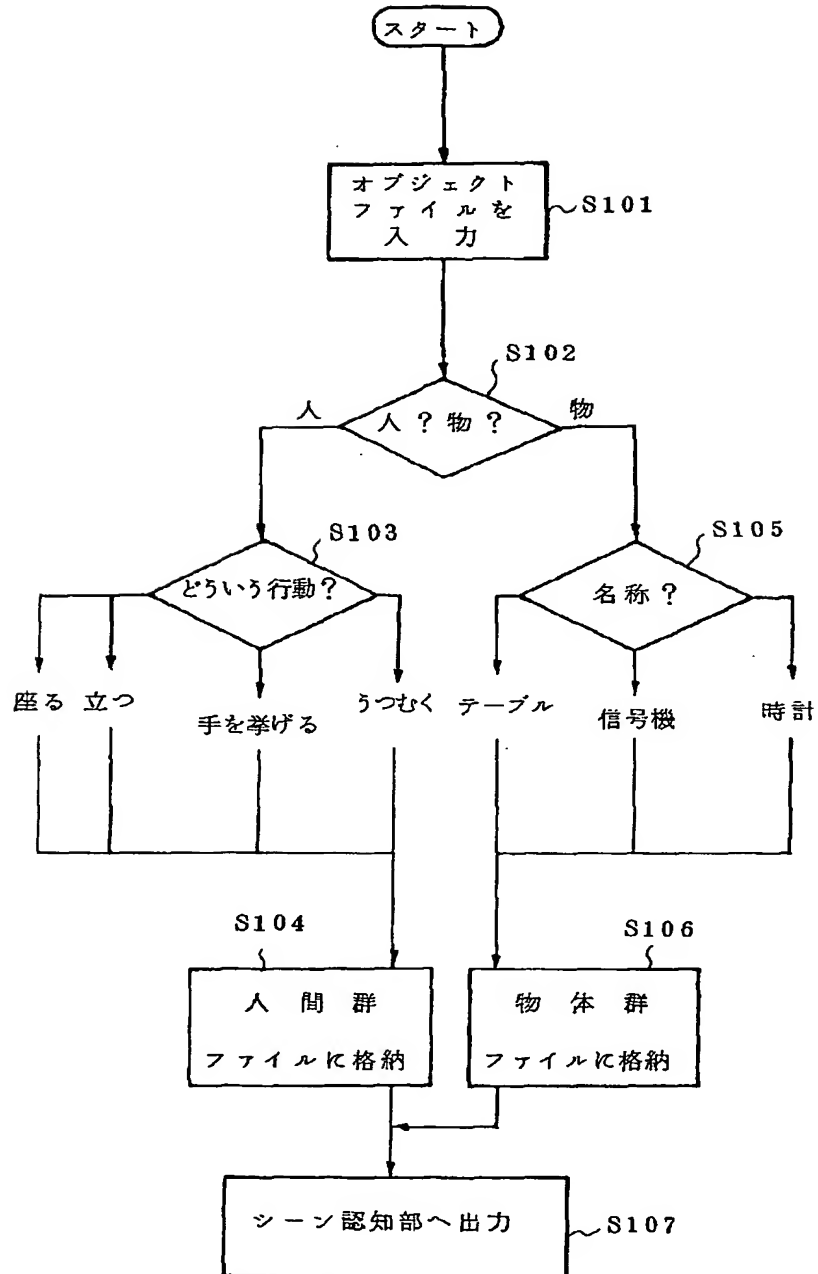


【図8】

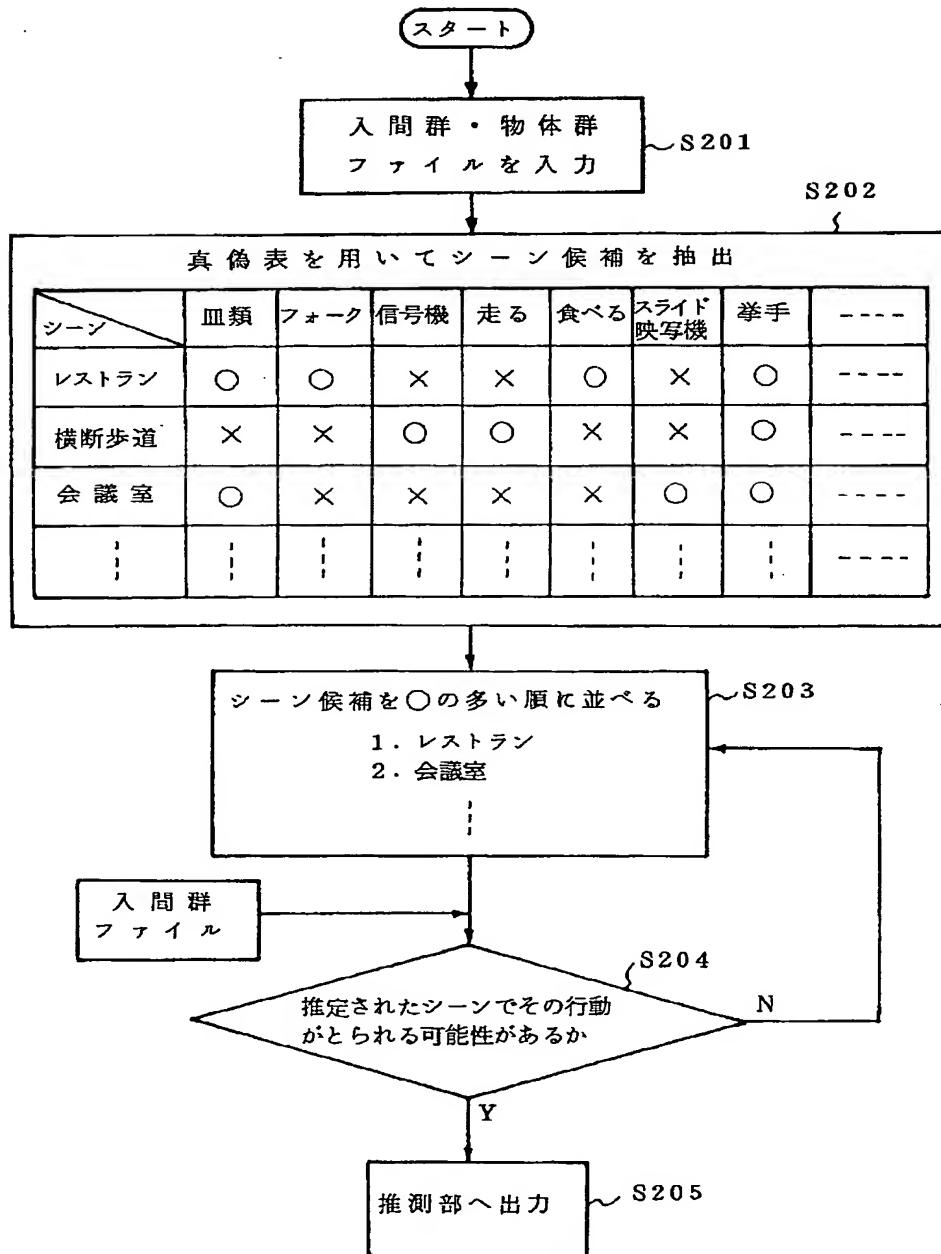




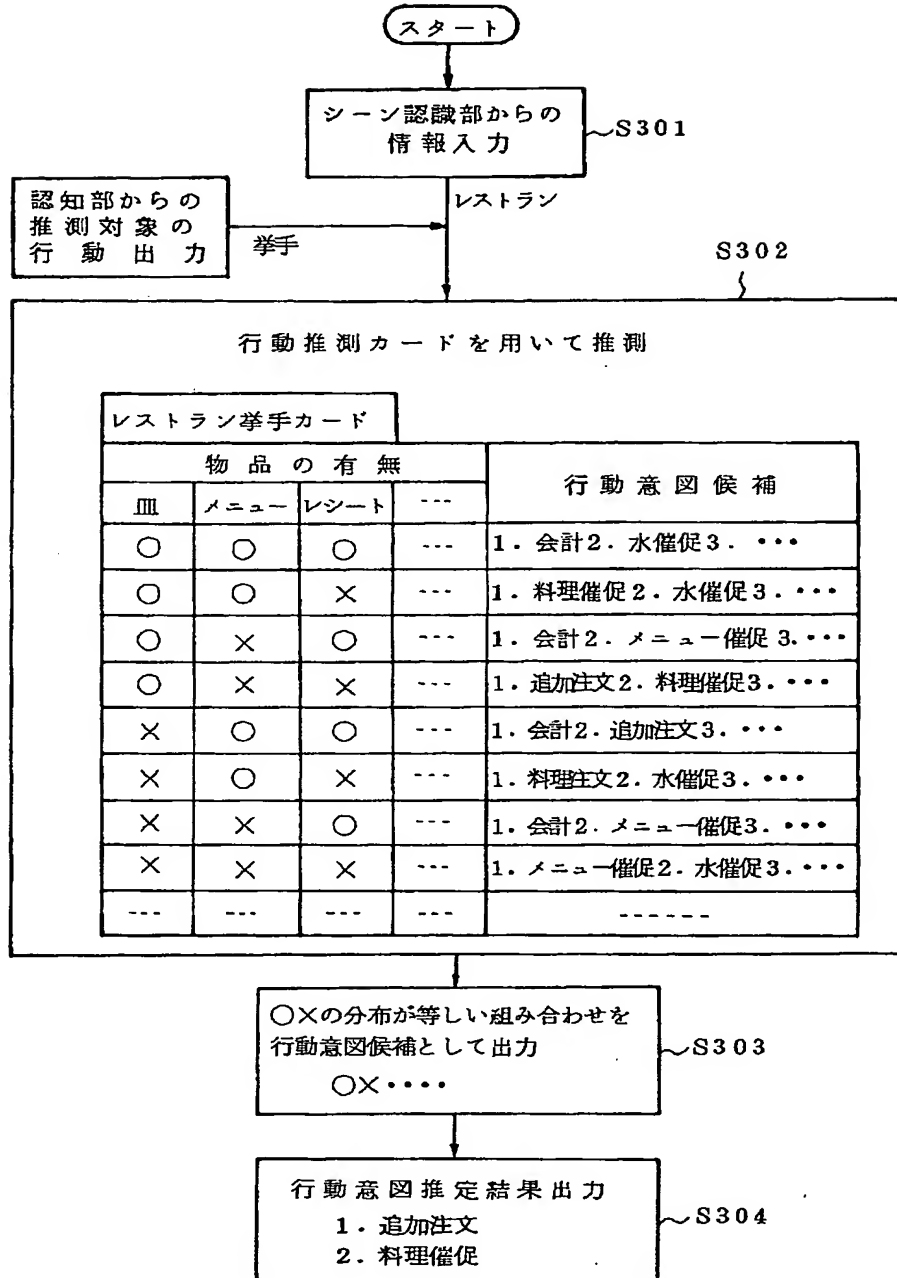
【図2】



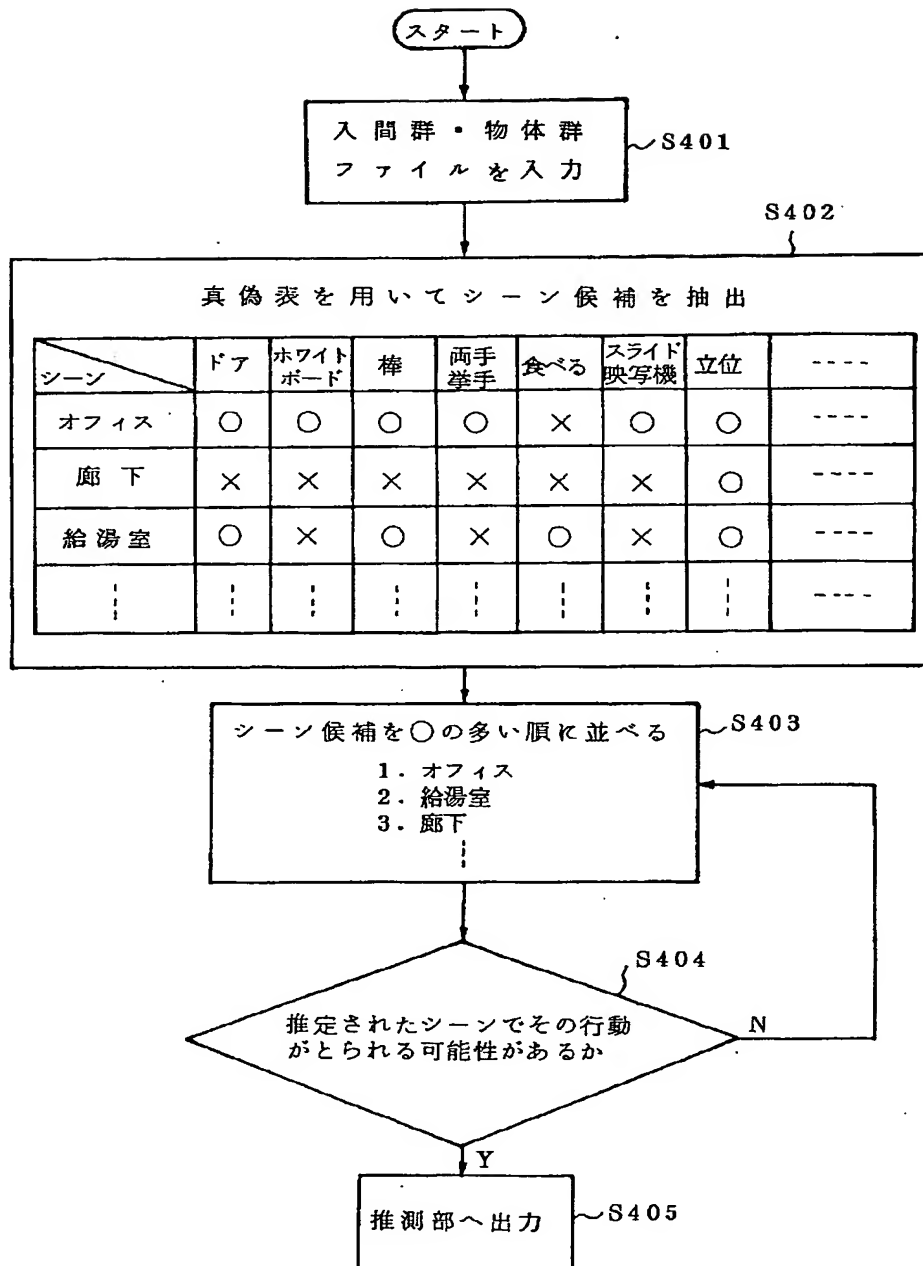
【図3】



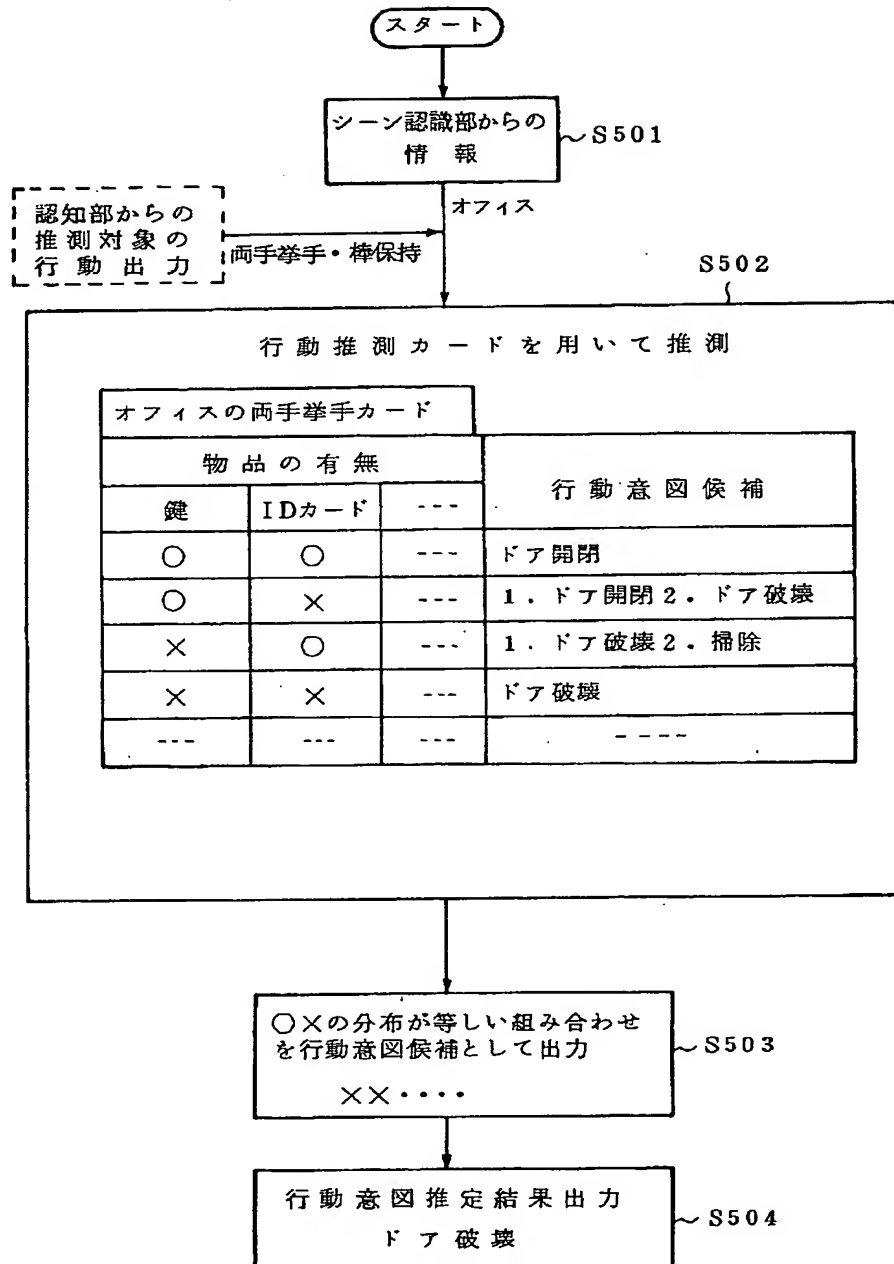
【図4】



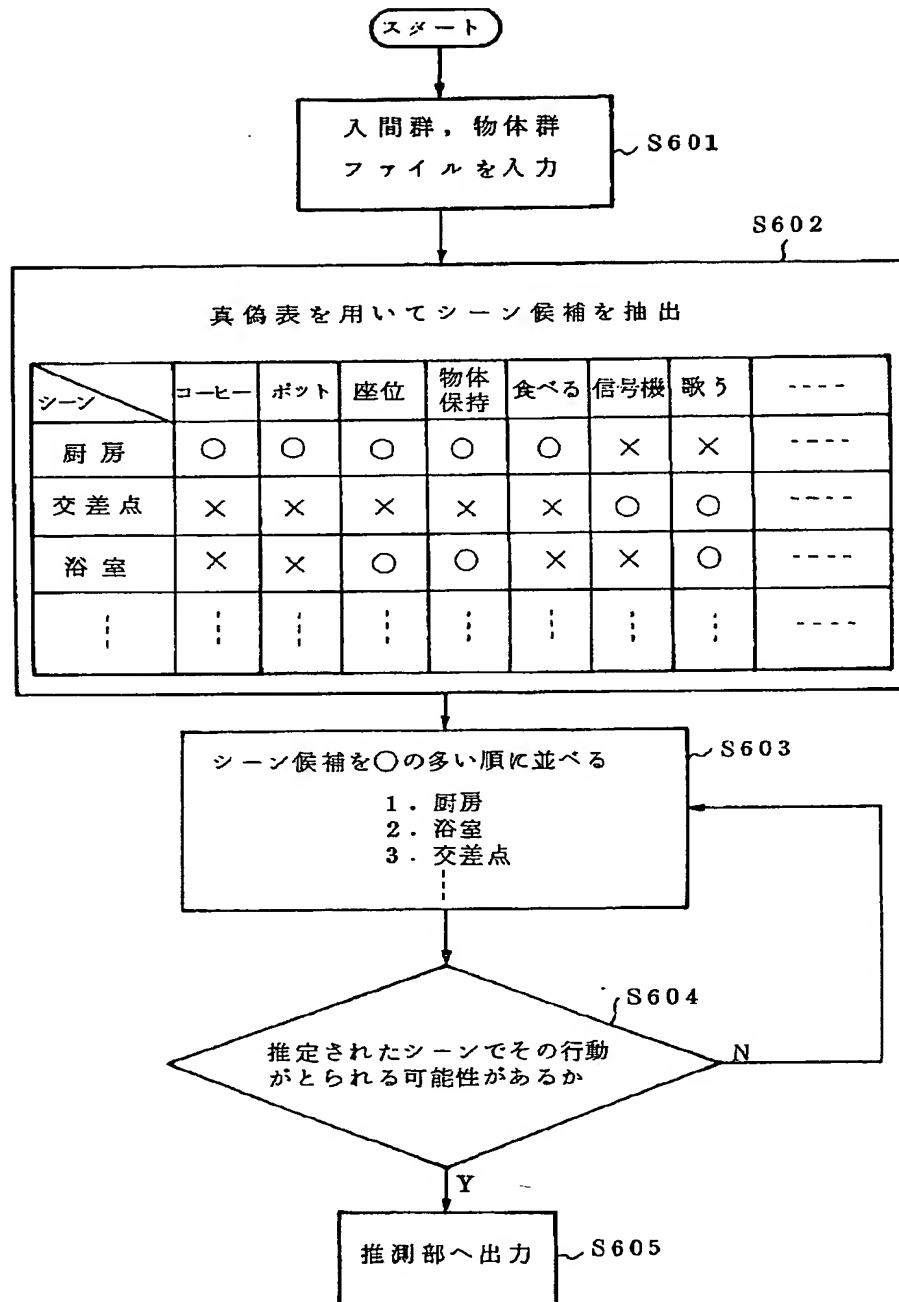
【図5】



【図6】



【図7】



フロントページの続き

(72)発明者 中村 信次  
愛知県豊田市トヨタ町1番地 トヨタ自動車株式会社内

(72)発明者 原 佳子  
愛知県豊田市トヨタ町1番地 トヨタ自動車株式会社内

Fターム(参考) 5L096 BA02 CA02 DA05 FA19 FA72  
HA09 HA13 JA16 KA17